Legibility and visual effectiveness of some proposed and current health warnings on cigarette packages

Health Canada - UPEI
Legibility Testing Laboratory

LEGIBILITY AND VISUAL EFFECTIVENESS OF SOME PROPOSED AND CURRENT HEALTH WARNINGS ON CIGARETTE PACKAGES

(Report of research done under contract to the Bureau of Tobacco Control, Health Protection Branch, Environmental Health Division, Health Canada, September, 1999)

Thomy Nilsson
Professor of Psychology
University of Prince Edward Island

LEGIBILITY AND VISUAL EFFECTIVENESS OF SOME PROPOSED AND CURRENT HEALTH WARNINGS ON CIGARETTE PACKAGES

(Report of research done under contract to the Bureau of Tobacco Control of Health Canada, September, 1999)

Thomy Nilsson
Professor of Psychology
University of Prince Edward Island

SUMMARY

At the Health Canada - UPEI Legibility Testing Laboratory, seven new designs for warnings on cigarette packages were tested along with two designs presently in use. Legibility was measured in terms of the maximal distance at which the warnings could be read. Visual effectiveness was measured using a rating scale. The new designs included pictures as well as words. Three sizes were evaluated. The results based on over 7,000 observations made by 14 persons with normal acuity and normal color vision had a reliability of 98% for legibility and 97% for effectiveness. The best new designs were about 2 times as legible and 3.5 times as effective as those in present use. Size of the printed words was the principle factor determining legibility. Doubling the size of the letters more than doubled the legibility. Warnings with bigger pictures were more effective than those with smaller pictures. Warnings with color pictures were more effective than those with black and white pictures. Some other implications of the data for the design of warnings are discussed. Establishment of a comparative standard to ensure legibility of warnings is recommended.

INTRODUCTION

In this study warnings that were read at a greater distance were defined as being more legible than warnings that had to be brought closer. An early use of distance to measure legibility of words printed in color was made by Preston, Schwankl & Tinker in 1932. Their results showed several color combinations that were more legible than black-and-white. However, the distance method was rarely used again, presumably due to the special facilities required by this method. Most subsequent research measured legibility in terms of the time required to read a word or message - words that could be read in briefer presentations were defined as more legible than words that required longer presentations. When asked by Health Canada to measure the legibility of warnings in color printed on colored backgrounds, Nilsson & Percival (1989) reasoned that neither briefly flashed words nor reading speed were representative of

what consumers did in a store when considering the purchase of cigarettes. In a feasibility study, they tried measuring legibility in terms of distance and found that it produced results that were both reliable and consistent with subjective impressions of reading difficulty. A contract to build an automated, 8 meter, test track led to establishment of the Health Canada - UPEI Legibility Testing Laboratory. A complete study in this facility showed that the color combinations used by several manufacturers for warnings on cigarette packages were substantially less legible than warnings printed in black-on-white (Nilsson, 1991). Under certain common lighting conditions the warnings on some brands were not legible at any distance. Similar problems were also found to exist with warnings on some over-the-counter medications.

Subsequent work for the Canadian Space Agency provided additional equipment[1] for the laboratory which enabled systematic study of how color affected the legibility of words and other visual information. The research with colored words on colored backgrounds revealed several color combinations that were more legible than black/white (Clements-Smith, Nilsson, Connolly, & Ireland, 1993). This was contrary to all existing guidelines at that time (Boff & Lincoln, 1988, Sanders & McCormick, 1993; Tinker, 1963). Since 1932, most research on the legibility of colored print had found that color per se had no effect on legibility - only the lightness contrast in the colors mattered (Knoblauch, Arditi & Szlyk, 1991; Legge, Parish, Luebker & Wurm, 1990). The discrepancy between our results and those of other researchers became understandable when we considered how color information is transmitted from the eye to the brain (Nilsson & Connolly, 1997): Color is carried by small neurons which can be densely packed in the retina for maximal image resolution. Brightness information is carried by large neurons to transmit major properties of the retinal image as fast as possible. Because transmission speed of neural fibres is proportional to their diameter, color information takes longer to reach the brain. Consequently when a visual task is defined in terms of speed, the brain will decide on the basis of the brightness information it receives first, and the color information is not used.

Additional research for the Canadian Space Agency demonstrated that the distance method could also be used to measure the effectiveness of graphic symbols, line drawings, image enhancement techniques, and commercial designs for potato packages (Clements-Smith, et al 1993). A recent project for NATO provided opportunity to compare recognition distance measurements with search-time measurements of the effectiveness of camouflage (Nilsson, 1999; Toet, Bijl, Kooi, & Valeton, 1998). Data from six UPEI students using distance correlated r = +.85 with the search time results of 60 NATO trained observers. Furthermore, the distance

---

[1] This equipment expanded the types of materials and presentation formats that could be tested. It included a standard light source that simulates daylight, and several other lighting conditions, the development of a display box that could accommodate 60 X 60 cm signs, and a unique viewing system that extends the effective length of the test track to 42 meters.

measurements discriminated the effectiveness of more camouflaged vehicles better than did search time.

Several studies conducted at the Health Canada - UPEI Legibility Testing Laboratory have demonstrated that the effectiveness of many types of graphic designs can be quantitatively measured in terms of the distance at which human observers can read or otherwise understand the information such designs are intended to convey whether printed, projected or presented on a computer monitor. It appears to be the only quantitative method which is sensitive to both the brightness and color characteristics of graphics. Furthermore it is the only method that produces data which can be directly related to the 20/20 distance system used to specify visual acuity for medical and legal purposes. For these reasons it was the method chosen to evaluate the effectiveness of new designs for warnings on cigarette packages, to compare those designs with each other, and to compare them with the warnings presently in use.

METHOD

Subjects

Fourteen undergraduate students were recruited in the UPEI Psychology Department to serve as subjects and paid for their time. The students were screened for normal acuity using a Snellen chart and for normal color vision using the Dvorine Test. The purpose was explained as involving measurements of the effectiveness of several designs and sizes of those designs for possible new warnings on cigarette packages, along with some designs presently being used. The measurement procedure was explained and demonstrated with a set of ten practice measurements. The testing procedure itself was under the subjects' control, and they could rest when they wanted.

Materials

Seven designs for new warnings printed on the front of cigarette packages were provided by Health Canada. These designs included pictures that illustrated the written warning. Three sizes which occupied 60%, 50%, and 40% of the front of the package were tested. The remainder of the front of the package was printed in a plain, medium-blue color. Two sets of one warning design, "Children See; Children Do", were distributed across the testing sequence. This duplication enabled determining the reliability of the measurement procedure. Also tested one-third and two-thirds through the sequence were two designs currently in use on cigarette packages. One of these had the warning printed in black-on-white; the other had the same warning in white-on-black. Figure 1 at the end of this report shows the seven new designs in their most legible size along with one of the currently used designs arranged left to right in order of decreasing legibility.

The printed words and graphics tended to be proportionally scaled within the warning area, but there were notable exceptions: 1) On both sets of the "Children see - ", on the "Cigarettes cause mouth cancer", and on the "Die hard smokers - " the words printed on the smallest designs were slightly bigger than the words on the medium size

design. 2) On the "Your Children are sick - " design, the words on the medium design were slightly bigger than those on the large design. 3) The two sets of the "Children see - " designs differed slightly in print quality. 4) The pictures in four of the designs, "Smoking kills babies", " -- mouth cancer", "Your children are sick --" and "This year --", were the same scale for all three warning areas. Only the cropping differed. 5) The 40% "Children see --" warnings had larger scale pictures than did the 50% warnings.

Apparatus

Printed designs for warnings on the front of cigarette packages were mounted inside a viewing box and illuminated from 45 degrees on each side by a pair of 100 watt General Electric "Soft White Delux" incandescent bulbs, which produced a luminance of 110 cd/m^2 and color temperature of 2700^0 K on a white Leeds-Northrup test plate. Baffled projection tubes on the lamp housings and black satin lining of the box reduced stray light inside the box to below measurable or visible levels. The box was mounted on a raised, 8 meter, linear-bearing track. A computer-controlled stepping motor moved the box with an acceleration and deceleration of 10 cm/sec^2 to and from a speed of 14 cm/sec. The subject was seated at one end of the track with his or her head position maintained by a chin rest.

Procedure

A back-and-forth procedure called the Psychophysical Method of Limits was used to measure the maximal distance at which the warnings could be read. Testing began with the warning inside the viewing box close to the subject - about 70 cm. When the subject started the measurement process by pressing a button, the warning moved gradually away. The subject was asked to press another button when she or he could no longer clearly read the warning. At that moment distance was measured even as the warning continued to travel further by a fixed plus a random distance before coming to a stop. When the subject next pressed the start button, the warning moved towards the subject. The subject was now asked to respond when the warning first became readable. These pairs of measurements were repeated five times and then the warning was changed. A third button enabled the subject to change his/her mind at any time until the next measurement was started. This repeated the previous measurement procedure. While rarely used, this option reduced the stress of the observing task.

The 26 designs were arranged in a sequence so that successive packages differed in both design and size. Designs with predominantly verbal messages were alternated with designs having colored pictures. All 26 designs were tested sequentially in a single, two-hour session for each subject. In a second session on another day, the measurements were repeated in reverse sequence to counterbalance order effects. Afterwards the subjects were shown each warning in succession and asked to rate the effectiveness of each using a 10 point scale. On this scale "0" represented "not effective"; "4 - 5" was "moderately effective"; and "9" was "very effective".

RESULTS

Ten running-averages of the five pairs of successive back-and-forth distances were calculated. The two averages that differed most from the mean were omitted. The mean and standard deviation of the remaining eight comprised one set of measurements. Since the warnings were tested twice in opposite order for each subject, there were 2 X 14 sets of 8 measurements for each design. The grand mean distances and standard error of the means are presented in Table 1 arranged in descending order of distance. (Longer distances indicate warnings that are easier to read.) The reliability of these measurements was checked by comparing the distances of the two identical sets of "Children see; children do" warnings that were distributed across the measurement sequence. The mean difference was 7 cm, which is 2% of their mean distance. This indicates the measurements were 98% reliable. The average standard error of the means of the 14 sets of measurements for each warning was 9.5 cm or 3% of their average distance. Since that variability includes general subject differences and order differences, the discriminability of the means was comparable to their reliability.

The "relative legibility" data in the right-hand column of Table 1 indicates the relative ease of reading a warning compared to the warning that could be read at the furthest distance. It is based on the inverse-square relationship between the area of an object's image in the eye and the distance of the actual object. Nilsson (1991), Clements-Smith, et al (1993), and Nilsson (1999) have found that retinal area was more closely related to subjective difficulty of recognizing graphic stimuli than distance itself. Since retinal area is proportional to the number of visual pathways carrying information about the image, the need for larger retinal images reflects a need for more information. In terms of relative legibility, the best designs were three times as easy to read as the poorest designs, and twice as readable as the current warnings

The most legible warnings, "Cigarettes cause strokes" and "Smoking kills babies", could be still be read at 3.9 meters. The least legible, "This year the equivalent - ", could only be only be read at up to 2.2 meters. Reading distance correlated +0.96 with the size of the letters. For a given size print, color affected legibility. Black letters on white backgrounds were slightly, but consistently, more legible than white-on-black. White letters on grey or flesh colored backgrounds were less legible than white-on-black.

The ratings of visual effectiveness for the two sets of "children see" warnings were 97% reliable. Table 2 shows the ratings data arranged in order of descending visual effectiveness. It provides a different perspective on the results. The most visually effective warning was the one about "mouth cancer" despite its relatively low legibility. All warnings that used color pictures except one were more effective than the warnings that used black and white pictures. For all designs with pictures that increased in scale as the warning area increased, namely the " - cause strokes", " Die hard --", and "Children see --" warnings, effectiveness increased with picture size. Differences in the effectiveness of warnings whose pictures were the same scale may be attributable to the influence of cropping, and larger letters obscuring the pictures.

Table 1.
Average maximum distances at which the various warnings could be read, average ratings of effectiveness, and other characteristics - arranged in descending order of legibility.

WARNING	SIZE (% of front)	LETTER HEIGHT (mm)	LETTER / BACKGRND COLOR	PICTURE	DISTANCE(cm)		**RELATIVE LEGIBILITY (%)**	RATING	
					MEAN	SE		MEAN	SE
cause strokes	60	7	black / w	color	387	10.7	**100**	6.6	.39
kills babies	60	7.8	w / grey	color	385	12.4	**99**	6.4	.38
die hard	60	6	black / w	color	360	7.5	**87**	6.4	.41
kills babies	50	7	w / grey	color	348	10.8	**81**	6.1	.40
children sick	50	6	w / black	b & w	343	9.7	**79**	5.6	.44
children sick	60	6	w / black	b & w	343	12.2	**79**	5.4	.52
children see	60	6	w / black	b & w	342	8.2	**78**	5.1	.54
children see	60	6	w / black	b & w	335	13.9	**75**	5.4	.43
mouth cancer	60	6	w / flesh	color	334	8.6	**75**	7.1	.36
cause strokes	40	5	black / w	color	315	10.6	**66**	6.4	.43
children see	40	5	w / black	b & w	311	10.1	**65**	5.1	.36
children sick	40	5	w / black	b & w	305	8.6	**62**	5.5	.53
kills babies	40	5	w / grey	color	302	7.5	**61**	5.9	.37
children see	40	5	w / black	b & w	301	6.5	**61**	4.9	.41
mouth cancer	50	5	w / flesh	color	294	10.3	**58**	7.1	.29
die hard	40	4.8	black / w	color	292	10.1	**57**	5.9	.45
mouth cancer	40	5	w / flesh	color	287	8.9	**55**	6.9	.29
children see.	50	4.8	w / black	b & w	282	7.8	**53**	4.9	.41
current - duM	33	4	black / w	none	282	6.4	**53**	2.1	.32
children see	50	4.8	w / black	none	278	7.4	**52**	4.9	.36
current - duM	33	4.2	w / black	none	278	8.1	**52**	1.9	.44
cause strokes	50	4	black / w	color	277	8.2	**51**	5.9	.41
die hard	50	3.7	black / w	color	273	9.9	**50**	6.4	.30
yearly deaths	60	4	black / w	b & w	271	7.8	**49**	3.0	.67
yearly deaths	50	3	black / w	b & w	225	10.9	**34**	3.4	.56
yearly deaths	40	2.5	black / w	b & w	223	14.3	**33**	3.5	.71

Table 2. Average maximum distances at which the various warnings could be read, average ratings of effectiveness, and other characteristics - arranged in descending order of effectiveness.

WARNING	SIZE (% of front)	LETTER HEIGHT (mm)	LETTER / BACKGRND COLOR	PICTURE	DISTANCE (cm)		RELATIVE LEGIBILITY (%)	RATING	
					MEAN	SE		MEAN	SE
mouth cancer	60	6	w / flesh	color	334	8.6	75	**7.1**	.39
mouth cancer	60	5	w / flesh	color	294	10.3	58	**7.1**	.38
mouth cancer	40	5	w / flesh	color	287	8.9	55	**6.9**	.46
cause strokes	60	7	black / w	color	387	10.7	100	**6.6**	.40
kills babies	60	7.8	w / grey	color	385	12.4	99	**6.4**	.44
cause strokes	40	5	black / w	color	315	10.6	66	**6.4**	.52
die hard	60	6	black / w	color	360	7.5	87	**6.4**	.54
die hard	50	3.7	black / w	color	273	9.9	50	**6.4**	.43
kills babies	50	7	w / grey	color	348	10.8	81	**6,1**	.36
kills babies	40	5	w / grey	color	302	7.5	61	**5.9**	.43
cause strokes	50	4	black / w	color	277	8.2	51	**5.9**	.36
die hard	40	4.8	black / w	color	292	10.1	57	**5.9**	.53
children sick	50	6	w / black	b & w	343	9.7	79	**5.6**	.37
children sick	40	5	w / black	b & w	305	8.6	62	**5.5**	.41
children sick	60	6	w / black	b & w	343	12.2	79	**5.4**	.29
children see	60	6	w / black	b & w	335	13.9	75	**5.4**	.45
children see	60	6	w / black	b & w	342	8.2	78	**5.1**	.29
children see	40	5	w / black	b & w	311	10.1	65	**5.1**	.41
children see	50	4.8	w / black	b & w	278	7.4	52	**4.9**	.32
children see	50	4.8	w / black	b & w	282	7.8	53	**4.9**	.36
children see	40	5	w / black	b & w	301	6.5	61	**4.9**	.44
yearly deaths	40	2.5	black / w	b & w	223	14.3	33	**3.5**	.41
yearly deaths	50	3	black / w	b & w	225	10.9	34	**3.4**	.30
yearly deaths	60	4	black / w	b & w	271	7.8	49	**3.0**	.67
current - duM.	33	4	black / w	none	282	6.4	53	**2.1**	.56
current - duM.	33	4.2	w / black	none	278	8.1	52	**1.9**	.71

DISCUSSION

Legibility

The least legible warnings were the "yearly deaths" warnings, which had the smallest print. The 40% size with 2.5 mm print had a maximum mean legibility distance of 223 cm. Based on the normal near-focus distance of 40 cm, this warning should be readable by people whose vision is as poor as 120/20.

Most of the designs that were tested used either black letters on white backgrounds or white letters on black backgrounds. The small but consistent advantage of black letters on white was also found by Clements-Smith, et al (1993). Sanders and McCormick (1993) suggest that white letters on black backgrounds should employ slightly narrower stroke widths than those for black-on-white to optimize legibility.

The use of colors other than black and white for words and backgrounds should also be considered as a means of making warnings more legible. Clements-Smith, et al (1993) found a number of color combinations that were significantly more effective than black-on-white. For words, those combinations were green/white, black/yellow, and green/yellow, with white/green, blue/yellow, black/red, and blue/white being at least equally effective, For solid symbols, black/red and black/yellow were significantly more effective than black/white; with blue/yellow, red/black, green/yellow, yellow/green, black/green, and blue/white being at least equally effective. For fine-line drawings, white/blue and white/black were significantly more effective than black/white, while yellow on black, yellow/blue, white/green, yellow/green, and red/black were equally effective. There is an evident change in the most effective color combinations as the width of the colored subject matter decreases. This suggests that the color combinations that are most legible for large letters may differ from those for small letters.[2]

The warnings with white letters on grey ("Smoking kills babies") and warnings white letters on flesh colors ("Cigarettes cause mouth cancer") were considerably less legible than warnings with the words on black or white backgrounds. Separating the words from the picture would improve legibility. However, superimposing words on pictures enables the use of bigger pictures, which improves visual effectiveness. It may be possible to use superimposed words with less loss of legibility if they are printed in certain colors. For example, while black letters on the "mouth cancer" warning would get lost in the darker areas of the picture, blue or orange letters might work better than either white or black. Similarly, yellow letters might be more effective on the grey smoke background of the "kills babies" warning. Testing the legibility of colored words and backgrounds in intermediate hues and a couple of levels of saturation may help to provide some general guidelines for the use of words together with colored pictures.

---

[2] I am preparing to systematically measure the effects of stroke width on the legibility of letters in all combinations of the six primary colors.

The words in fine print providing additional information in the warnings were not tested because the warnings could not be brought closer than 80 cm to the subject with the present apparatus. Ninety centimetres was approximately the maximum distance at which the 2 mm, fine print in the "yearly deaths" warnings could be read by people with normal, 20/20 vision. Since the average person has a near focus of about 40 cm, non-nearsighted people with poorer than 20/40 vision would have difficulty reading this information.[3] Such testing will require a special viewing box to obtain the same glare free, uniform illumination with minimal scattered light. The size of this fine print is similar to the size used for warnings and instructions on the packages of many medications sold over the counter. Therefore, an ability to test the legibility of small print would enable the Health Canada - UPEI Legibility Laboratory to measure the legibility of various types of health information on consumer products.[4]

All the present warnings had similar, moderately glossy surfaces and did not involve glossy or metallic inks. Nilsson (1991) found that specular reflectance such as produced by a glossy coating or metallic inks could severely degrade legibility in certain common illumination conditions. It is reasonable to expect that similar degradation would occur if a glossier coatings or inks had been used in the present warnings. The warnings would be legible under a wider range of lighting conditions if they were provided with a flat or matte surface finish.[5]

Visual Effectiveness

The ratings of visual effectiveness correlated only +0.57 with legibility. In other words, less than one-third of the variance of their respective mean values was common to both sets of measurements. Given the high, 97% and 98%, reliability of both sets, this modest commonality indicates that legibility and visual effectiveness were largely independent effects in the warnings' designs. This is illustrated by the following. While the three warnings with the highest legibility ("cause strokes", "kills babies", "die hard", all in the large, 60% size) were also quite effective, the three most effective warnings were all three sizes of the "mouth cancer" warnings. The lower standard error of their ratings compared to the former's indicates substantial subject agreement on this preference.

Several factors may contribute to the effectiveness of the mouth cancer warnings: (1) the alarming nature of the picture, (2) the pictures' large size, (3) the

---

[3] Some people who are nearsighted can focus images considerably closer than 40 cm, but there are other reasons why visual acuity may be poorer than 20/40.

[4] A close-up viewing box for testing fine print will be developed for the Health Canada - UPEI Legibility Laboratory come the end of classes this academic year.

[5] Readers can see this for themselves by covering one of the warnings in Figure 1 with a flat-finish cellophane mending tape. Bearing in mind that the tested package fronts were somewhat more glossy that the copies in Figure 1, note the reduction in glare from the flatter surface as the warning is tilted in proximity to a lamp.

implication for personal appearance. Persons the age of our undergraduate students may be more sensitive to the consequences of smoking on appearance than to consequences which relate to invisible medical problems, medical problems that take years to develop, or to young children. This suggests that warnings on cigarette packages might have more impact on particular groups if designed specifically to address the concerns of those groups. Further testing with subjects from various groups would be required to determine how various design themes are perceived by adolescents, young adults, young parents, older adults, etc.

As a whole, warnings with color pictures were judged more visually effective than those with black and white pictures. However, since the colored and black-and-white warnings also involved different pictures and themes, the advantage of color is not unequivocally proved by the present study. For similar reasons it is not possible to ascertain how full-size pictures with overlaid words ("mouth cancer" and "kills babies") compare in effectiveness to those where the pictures and words are separated for better legibility ("cause strokes", "die hard"). As mentioned in the legibility section, the existence of color combinations that are more or equally as legible as black-on-white suggests that it may be possible to develop warnings with full size color pictures and overlaid words that would maximize both visual effectiveness and legibility.

Standards for Legibility and Visual Effectiveness

With hundreds of letter fonts, millions of color combinations, and innumerable other options for pictures, it is evident that guidelines would not be adequate to ensure acceptable warnings. There is an alternative to the use of guidelines and specifications. It involves a standardized, simple method of measuring legibility and visual effectiveness in terms of distance along with a "legibility standard".

Legal requirements for vision are already based on distance. The 20/50 criterion of adequate vision for driving means that the driver must be able to see at 20 feet what the average person can see at 50. Distance specifications of legibility and effectiveness can therefore be directly related to medical and legal practices. Consider the following. For a person with 20/20 vision, comparing a warning with a legibility distance of 360 cm with a warning that has a legibility distance of 180 cm is like degrading that person's vision to 20/40.[6] Furthermore, a warning with a legibility distance of 80 cm for people with normal vision would be unrecognizable at any distance by people with 20/60 vision because they could not bring the warning close enough before it was too close to be focused. Since the ability to focus closely deteriorates with age, this limitation should be considered in the design of warnings intended to be readable for elderly people.

The distance method is also conceptually simple to implement. In principle it only requires a tape measure and a certain level of even, glare-free illumination from a

---

[6] A reader with 20/20 vision or vision corrected to 20/20 may experience this effect by borrowing a pair of glasses from a person with 20/40 vision. Compare any warning with how it appears with and without those extra glasses illustrates the lose of legibility that results when legibility distance is halved.

continuous spectrum source such as a light bulb. No complex electronics or optics are essential. Guidelines could be provided for the design of a basic system that would enable manufacturers and design houses to make their own arrangements for evaluating warnings during development.

A "legibility standard" need only be a certain, agreed upon warning, such as one of the proposed warnings. The acceptability of any other warning would be determined by comparing its relative legibility to that of a copy of the standard using the same measuring facility for both designs. With the standard's relative legibility set to 100%, any design regardless of color, font, etc. whose relative legibility was at least 75% (for example) could be defined as acceptable. The provision of a legibility standard reduces the need for stringent specification of the testing apparatus, since variations that might somewhat increase or decrease the legibility distance of the design being tested would similarly affect the measurements of the standard. The distance method together with a legibility standard offers a robust means of ensuring visually adequate warnings.[7]

---

[7] I would be pleased to further discuss the standardization issue and how the distance method could be implemented. - TN

RECOMMENDATIONS

Based on the results of the present study and related research, the following are recommendations that would help maximize the legibility and visual effectiveness of health warnings:

1. Use the largest letters possible. Letter size was more important than size of the warning area, though bigger letters of course require more space.

2. Place the words on a uniform background instead of on top of pictorial details.

3. Black letters on white backgrounds are slightly more legible than the reverse. Whether this holds for small print remains to be determined. Sanders and McCormick (1993) suggest that the legibility of white letters on black may be optimized by using slightly narrower stoke widths than those for black/white.

4. When other aspects of the warning design lead to a background color that is neither black or white, a letter color which is neither white nor black may maximize legibility. Some guidelines for the six primary colors are presented in the Discussion. Further systematic testing of additional color combinations would assist designers.

5. The legibility of warnings can be improved by the use of letter/background colors other than black and white. Color is also widely noted for its ability to attract attention in graphics. Clements-Smith, et al (1993) cite some examples. The effects of letter size on the color combinations that are optimal should be determined.

6. Avoid glossy surface coatings and metallic inks. Under certain common lighting conditions, these produce glare which can severely degrade legibility (Nilsson, 1991). A flat or matte finish would make the warnings legible under a wider range of lighting conditions.

7. Use color pictures to improve visual effectiveness.

8. Use the biggest pictures possible to improve visual effectiveness.

9. The number of relevant factors such as color, size, interactions between words and pictures, style, and theme of both the words and picture in a warning make specifications an unfeasible means of ensuring legibility or visual effectiveness. Instead, a comparative standard for the legibility and effectiveness of health warnings should be established. The acceptability of any warning design could then be specified relative to this standard using distance measurements.

ACKNOWLEDGEMENTS

Donnay MacNalley helped prepare the test cards. She and Jill MacDonald served as subjects and tested most of the other subjects. I thank the following persons for their participation as subjects: Erin Clow, Tasha Gallant, Lisa Hogan, Mini Lakelle, Steve Long, Catherine MacDougall, Rachal MacLeod, Paige Murphy, Patrice Shea, Philip Sheffrington, Robert Smith, Dierdre Squire.

REFERENCES

Boff, K.R. and Lincoln, J.E. (1988) *Engineering Data Compendium: Human Perception and Performance*. Wright-Patterson Air Force Base, Ohio, 2272-2277.

Clements-Smith, G,, Nilsson, T.H., Connolly, G.K. and Ireland. W. (1993) *Development of an Industrial Service Utilizing New Technology to Optimize Visual Information Displays for Both Human and Machine Vision*. (Report 9F006-2-0021) Canadian Space Agency, Ottawa.

Knoblauch, K., Arditi, A. and Szlyk, J. (1991) Effects of chromatics and luminance contrast on reading. *Journal of the Optical Society of America A*, 8, pp. 428-439.

Legge, G.E., Parish, D.H., Luebker, A. and Wurm, L.H. (1990) Psychophysics of reading XI: Comparing color contrast and luminance contrasts. *Journal of the Optical Society of America A*, 3, pp. 400-51.

Nilsson, T.H. (1991) *Legibility of Tobacco Health Messages with Respect to Distance*. Report to the Tobacco Products Division of Health & Welfare Canada (contract H4078).

Nilsson, T.H. (1999) Evaluation of target acquisition difficulty using recognition distance to measure required retinal area. NATO SCI-12 Workshop on Search and Target Acquisition. Utrecht, The Netherlands.

Nilsson, T.H. and Connolly, G.K.(1997) Chromatic contribution to shape perception revealed in a non-temporal task using distance, in *Colour Vision Research: Selected Proceedings of the International Conference, John Dalton's Colour Vision Legacy*. C. Dickinson, I. Murphy and D. Carden, Eds., pp. 197-206, Taylor and Francis, Bristol, Pennsylvania, 1997.

Nilsson, T.H. and Percival, T.Q. (1989) *Evaluating the Legibility of Tobacco Health Warnings with a Method Suitable for Defining Market Guidelines*. (Report K302507), Health and Welfare Canada, Ottawa, 1989.

Preston, K., Swankl, H.P. and Tinker, M.A. (1932) The effect of variations in color of print and background on legibility", *Journal of General Psychology*, 6. pp. 459-461.

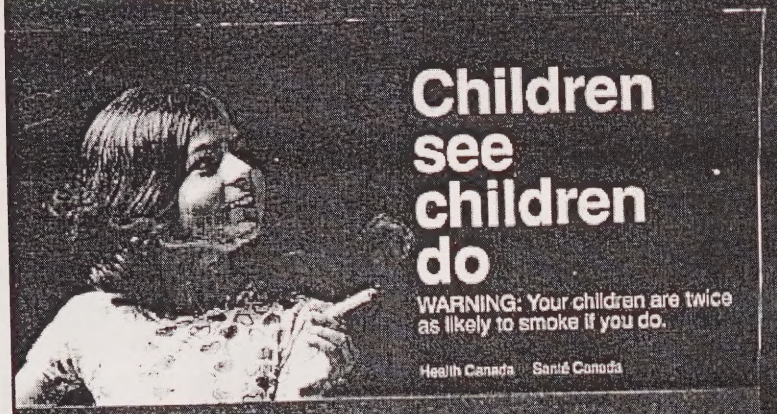
Sanders, M.S. and McCormick, E.J. (1993) *Human Factors in Engineering Design*. McGraw-Hill, New York.

Tinker, M.A. (1963) *Legibility of print*. Iowa State University Press, Ames, Iowa.

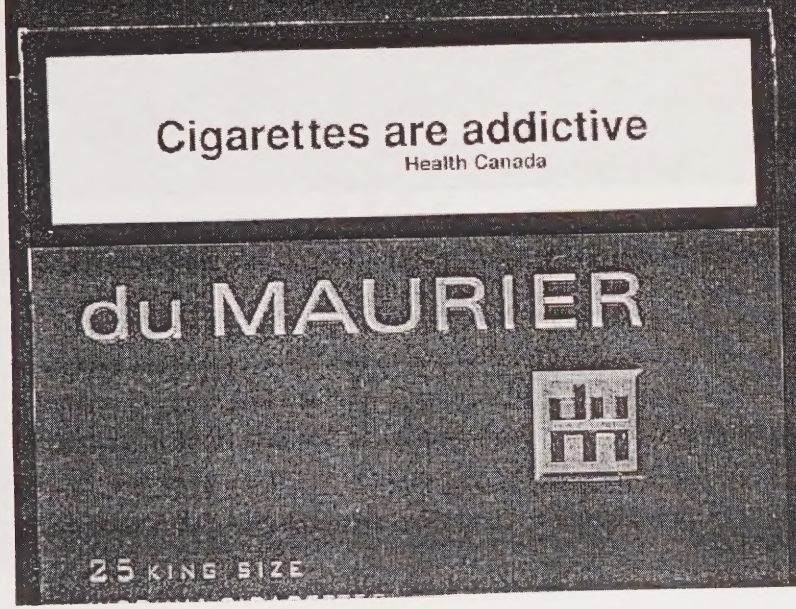
Toet, A., Bijl, P., Kooi, F.L. and Valeton, J.M., (1998) *A High-Resolution Data Set for Testing Search and Detection Models*. (Report TM-98-A020), TNO Human Factors Research Institute, Soesterberg, The Netherlands.

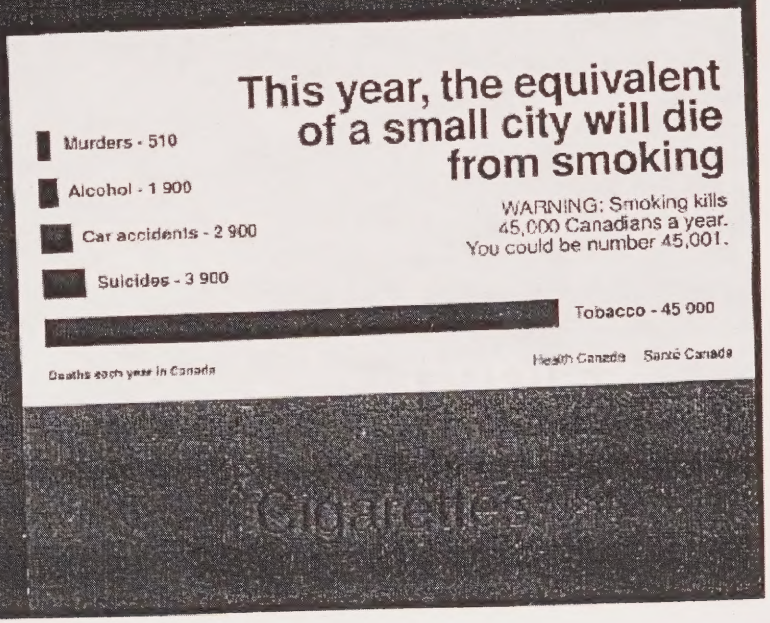
Figure 1. (on two following pages) The eight designs for warnings on cigarette packages that were tested for legibility and visual effectiveness. The designs are arranged from left to right in order of decreasing legibility. The most legible size of each design is shown.


Cigarettes cause strokes
burst blood vessel
cross section of brain
WARNING: Cigarette smoke clogged this smoker's arteries with plaque. The dark spot shows a burst blood vessel in the brain. This stroke caused death.
Health Canada Santé Canada
Cigarettes
Smoking kills babies
WARNING: Smoking near babies can cause Sudden Infant Death Syndrome.
Health Canada Santé Canada
Cigarettes
Die hard smokers often die hard deaths
healthy lung
diseased lung
WARNING: Cigarettes cause lung cancer. People with lung cancer often die hard, lingering and painful deaths.
Health Canada Santé Canada
Cigarettes
Your children are sick of your smoking
WARNING: Second-hand smoke triggers asthma attacks.
Health Canada Santé Canada
Cigarettes


Children see children do
WARNING: Your children are twice as likely to smoke if you do.
Health Canada Santé Canada
Cigarettes


Cigarettes cause mouth cancer
WARNING: Cigarettes make it harder for your saliva to remove germs in your mouth. Smoke also increases the risk of cancer, gum disease and tooth loss.
Health Canada Santé Canada
Cigarettes


Cigarettes are addictive
Health Canada
du MAURIER
25 KING SIZE


This year, the equivalent of a small city will die from smoking
Murders - 510
Alcohol - 1 900
Car accidents - 2 900
Suicides - 3 900
Tobacco - 45 000
WARNING: Smoking kills 45,000 Canadians a year. You could be number 45,001.
Deaths each year in Canada
Health Canada Santé Canada
Cigarettes